Gender-Based Analysis Plus

For Spouse or Common-Law Partner in Canada Class AA Pilot (Model)

Document Change Control

Version #	Changed by	Date of Change	Description of Changes
0.1	Solutions Task Team	May, 2021	First Draft
0.2	Advanced Analytics Solutions Centre	June, 2021	Added comments made by SPP-Digital Policy, SPP-GBA+ and IB.
1.0	Advanced Analytics Solutions Centre	July, 2021	Updated and modified according to the comments.
2.0	Advanced Analytics Solutions Centre	March, 2022	Follow-up analysis for monitoring

Table of Contents

1. Background

s.16(1)(b)
s.16(2)(c)

In response to the increased targets in the 2021 Levels plan, the Department has developed an Advanced Analytics (AA) model to assist in the processing of applications in the Spouse or Common-Law Partner in Canada Class. The model uses a variety of criteria to assess the eligibility of both the sponsor and the principal applicant (PA.) Based on its assessment the model assigns applications to either a "Green" or a "Standard" bin. Applications in the Green Bin receive a positive administrative eligibility determination based solely on the model's recommendation, while Standard Bin applications will be processed by officers as per normal procedures. For all applications, a human officer is responsible for assessing the eligibility of dependents, the admissibility of the PA and dependents, and for taking the final decision.

2. Expected Overall Impact

The model's overall impact on clients is expected to be minimal given that the model neither refuses applications nor makes any form of negative recommendation, but rather assigns applications into one of the two bins. Accordingly, the "worst" potential outcome for any application passed through the model would be that it is processed by an officer as per normal procedures.

3. Analytical Method

This analysis proceeds primarily by comparing, along different dimensions, the proportion of applications in an inventory before it is run through the model to the proportion of applications in the Green and Standard bins after assessment, and noting any significant variations, or lack thereof.

For example, if same-sex partnerships account for 20% of the inventory fed into the model, but then account for only 5% of the Green Bin, this could indicate that the model may have introduced an unintended bias against same-sex couples. Of course, this is not *necessarily* so. There may be something in the initial inventory that explains and justifies the discrepancy.

Furthermore, each triage criterion is explained and justified in the *Officer of Record* memo to assign accountability of model decisions and to prevent any unintended, differential distribution of triage results.

The initial analysis was conducted on the open applications that were triaged by the model for the first time, most of which were backlogs. The set comprised of approx. 6,000 open applications in the SCLPC inventory when the model was deployed initially. In June 2021, a minor update was made to the model to enhance the program integrity. In order to monitor the model impact to identify and mitigate any unintended outcomes, a follow-up analysis was conducted by following the same method in March 2022, when the number of triaged applications by the model reached approx. 34, 000.

4. GBA+ Data Tables

4.1 Model Development and Testing Data

- **Gender Dimensions of the Green versus Standard Bin**

The data set used to develop and test the AA model contained 40,343 applications, 23,555 (58.4%) of which had female PAs, while 16,785 (41.6%) had male PAs. When these applications were run through the model, it passed 23,942 into the Green Bin and 16,401 into the Standard Bin.

The gender profile of both bins closely mirrored that of the initial data set. In the Green Bin, 14,238 (59.5%) of the PAs were female, while 9,702 (40.5%) were male. Similarly, of the 16,401 applications passed into the Standard Bin, 56.8% of the PAs were female, while 43.2% were male.

Table 1.1: Testing Dataset by PA Gender

PA Gender	Total #	Total %
Female	23,555	58.4%
Male	16,785	41.6%
Other	3	-
All	**40,343**	**100.0%**

Table 1.2: Testing Dataset by Bin and PA Gender

Bin	PA Gender	Total #	Total %
Green	**Female**	14,238	59.5%
	Male	9,702	40.5%
	Other	2	-
	All	**23,942**	**59.4%**
Standard	**Female**	9,317	56.8%
	Male	7,083	43.2%
	Other	1	-
	All	**16,401**	**40.7%**

Relatedly, of the 40,343 applications in the development and testing inventory, the sponsor and the PA were of a different sex in 38,747 (96%) applications, while 1,596 (4%) were of the same gender. In the Green Bin, 23,054 (96.3%) relationships were opposite sex, while 888 (3.7%) were same-sex. Given that the gender profile of both the individual applicants and their relationships in the Green and Standard bins so closely matches the related profiles in the development and testing data set, the model did **not** result in any differential output along the dimension of gender.

Table 1.3: Testing Dataset by Same-Sex Couple

	Total #	Total %
Not Same Sex Couple	38,747	96.0%
Same Sex Couple	1,596	4.0%
All	**40,343**	**100.0%**

Table 1.4: Testing Dataset by Bin and Same-Sex Couple			
Bin	**Couple**	**Total #**	**Total %**
Green	**Not Same Sex Couple**	23,054	96.3%
	Same Sex Couple	888	3.7%
	All	**23,942**	**59.4%**
Standard	**Not Same Sex Couple**	15693	95.7%
	Same Sex Couple	708	4.3%
	All	**16,401**	**40.7%**

4.2 Model Deployment – Initial Analysis

- **Gender Dimensions of the Green versus Standard Bin**

The model was run over a set of open applications for the first time on April 27th, 2021. The set comprised 6,046 open applications in the SCLPC inventory. In 3,380 (55.9%) of the 6,046 applications the PA was female, while 2,666 (44.1%) had male PAs; 3,483 (57.6%) were assigned into the Green Bin and 2,563 (42.4%) into the Standard Bin.

Table 2.1: Applications That are Triaged in the Model's First Deployment by PA Gender		
PA Gender	**Total #**	**Total %**
Female	3,380	55.9%
Male	2,666	44.1%
Other	0	0.0%
All	**6,046**	**100.0%**

Table 2.2: Applications That are Triaged in the Model's First Deployment by Bin and PA Gender			
Bin	**PA Gender**	**Total #**	**Total %**
Green	**Female**	2,012	57.8%
	Male	1,471	42.2%
	Other	0	0.0%
	All	**3,483**	**57.6%**
Standard	**Female**	1,368	53.4%
	Male	1,195	46.6%
	Other	0	0.0%
	All	**2,563**	**42.4%**

The sponsor and PA were different-sex in 5,795 (95.8%) applications and same-sex in 251 (4.2%) applications.

Table 2.3: Applications That are Triaged in the Model's First Deployment by Bin and Same-Sex Couple								
	Total #	**Total%**	**GB#**	**GB%**	**GB Rate**	**SB#**	**SB%**	**SB Rate**
Not Same Sex Couple	5,795	95.8%	3,334	95.7%	57.5%	2,461	96.0%	42.5%
Same Sex Couple	251	4.2%	149	4.3%	59.4%	102	4.0%	40.6%
Total	**6,046**	**100.0%**	**3,483**	**100.0%**	**-**	**2,563**	**100.0%**	-

This gender profile remained largely unchanged after this set was run through the model and applications assigned into the Green and Standard bins. Of the 3,483 applications passed into the Green Bin, the PA was female in 2012 (57.8%) applications and male in 1,471 (42.2%). Similarly, in the 3,483 applications in the Green Bin, 3,334 (95.7%) were for opposite-sex couples while 149 (4.3%) were for same-sex couples. A similar gender distribution is noted in the Standard Bin, where the PA was female in 1,368 (53.4%) applications and male in 1,195 (46.6%) applications.

As was the case with the developmental data set, therefore, the gender profile of the initial inventory is essentially maintained in both the Green and Standard bins. As such, it again does not appear that the model has introduced any differential outputs along gender lines in its initial deployment.

4.3 Model Deployment – Follow-up Analysis for Monitoring

- **Gender Dimensions of the Green versus Standard Bin**

By mid-March 2022, the model has triaged 34,018 SCPLC applications, of which 18,418 (54.1%) applications had female PAs, while 15,574 (45.8%) had male PAs. Among these triaged applications, 17,640 were assigned into the Green Bin and 16,378 into the Standard Bin.

Table 3.1: Applications That are Triaged by the Model by PA Gender- Follow-up Analysis

PA Gender	Total #	Total %
Female	18,418	54.1%
Male	15,574	45.8%
Other	7	0.0%
Unknown	19	0.1%
All	**34,018**	**100.0%**

Table 3.2: Applications That are Triaged by the Model by Bin and PA Gender- Follow-up Analysis

Bin	PA Gender	Total #	Total %
Green	**Female**	9,955	56.4%
	Male	7680	43.5%
	Other	1	0.0%
	Unknown	4	0.0%
	All	**17,640**	**51.9%**
Standard	**Female**	8,463	53.4%
	Male	7,894	46.6%
	Other	6	0.0%
	Unknown	15	0.1%
	All	**16,378**	**48.1%**

The sponsor and PA were different-sex in 32,572 (95.7%) applications and same-sex in 1,456 (4.3%) applications.

Table 3.3: Applications That are Triaged by the Model by Bin and Same-Sex Couple – Follow-up Analysis

	Total#	Total%	GB#	GB%	GB Rate	SB#	GB%	SB Rate
Not Same Sex Couple	32,562	95.7%	16,878	95.7%	51.8%	15,684	95.8%	48.2%
Same Sex Couple	1,456	4.3%	762	4.3%	52.3%	694	4.2%	47.7%
Total	**34,018**	**100%**	**17,640**	**100%**	-	**16,378**	**100%**	-

This gender profile remained largely unchanged with the ongoing implementation of the model since its initial deployment, and applications assigned into the Green and Standard bins. Of the 17,640 applications passed into the Green Bin, the PA was female in 9,955 (56.4%) applications and male in 15,574 (45.8%). Similarly, in the 17,640 applications in the Green Bin, 16,878 (95.7%) were for opposite-sex couples while 762 (4.3%) were for same-sex couples, and the percentages remained same as the results of the analysis for the first fun of the model. A similar gender distribution is noted in the Standard Bin, where the PA was female in 8,463 (53.4%) applications and male in 7,894 (46.6%) applications. In the 16,378 applications in the Standard Bin, 15,684 (95.8%) were for opposite-sex couples while 694 (4.2%) were for same-sex couples.

The gender profile of the triaged applications in the follow-up analysis is essentially maintained in both the Green and Standard bins. As such, it again does not appear that the model has introduced any differential outputs along gender lines in since its launch.

s.16(1)(b)

s.16(2)(c)

5. Summary

Based on the initial and follow-up analysis, data from testing and implementation stage indicate that the model does not currently appear to have any significant, disproportionate impact on clients. Assignments into the Green and Standard bins appear relatively equitable, with notable variances easily understood following review of the data.

In terms of gender profile, both the individual applicants and their relationships in the Green and Standard bins closely matches the related profiles in the testing dataset and model implementation for both initial and follow-up analysis, thus the model did not result in any differential output along the dimension of gender.

An examination of country-of-birth data revealed that the model eligibility approval rate is being evenly 'distributed' to most countries. In the case of a few countries, where more notable variations were found, a closer examination of the data showed that applications from those countries more commonly failed to meet certain model rules. As mentioned in the earlier section of this analysis, the model's overall impact on clients is expected to be minimal because the model neither refuses applications nor makes any refusal recommendations, but rather assigns applications into one of the two bins. Therefore, the applications that were triaged in the standard bin with PAs born in these countries are processed by an officer as per normal procedures, which may still result with an approved final decision.

It is observed that no unintended bias is immediately apparent in the model's initial output and ongoing deployment. The overall potential for any negative impacts on specific groups of applicants is mitigated in large part by the fact that the model only automates certain positive eligibility decisions and that all final decisions are always made by an officer.

6. Annex: Additional Data Tables

s.16(1)(b)
s.16(2)(c)

s.16(1)(b)
s.16(2)(c)

Protected B and ATIP Exempt

s.16(1)(b)
s.16(2)(c)

Protected B and ATIP Exempt

s.16(1)(b)
s.16(2)(c)

Protected B and ATIP Exempt

s.16(1)(b)
s.16(2)(c)

7

s.16(1)(b)
s.16(2)(c)

s.16(1)(b)
s.16(2)(c)

Protected B and ATIP Exempt

s.16(1)(b)
s.16(2)(c)

Protected B and ATIP Exempt